

Db 421 GTPNTTAKFSDKLASTNSNEDNMFMKITNLKVGAAPDVRSTINUSPGVYTGAYQD 480  
 Qy 481 LLMFKDNSTNHILLEPELRNFEPPTKTSQSLTQEITISWSPNQARKSQEEYKKYFE 540  
 Db 481 LLMFKDNSTNHILLEPELRNFEPPTKTSQSLTQEITISWSPNQARKSQEEYKKYFE 540  
 Qy 541 GALGEDDNIDFAQNTVLKDQYSKKLSSMKTANKEITYTQLGDKISYEAASCNLSK 600  
 Db 541 GALGEDDNIDFAQNTVLKDQYSKKLSSMKTANKEITYTQLGDKISYEAASCNLSK 600  
 Qy 601 DPFSSILYQKNTIEGETAYYYVADAETKEIDKYRIPIQISNRNKLTFIGHGKSEFT 660  
 Db 601 DPFSSILYQKNTIEGETAYYYVADAETKEIDKYRIPIQISNRNKLTFIGHGKSEFT 660  
 Qy 661 DTFANLVDYDLSSEIETLNLLAKADISPKYIEINLGCMNPMSYISIAYETTPGKLLIK 720  
 Db 661 DTFANLVDYDLSSEIETLNLLAKADISPKYIEINLGCMNPMSYISIAYETTPGKLLIK 720  
 Qy 721 DRYSELMPSSISODSISITYSANQEVIRINEEGKREILDHSGKWNKKEESIKDISSEKEYSF 780  
 Db 721 DRYSELMPSSISODSISITYSANQEVIRINEEGKREILDHSGKWNKKEESIKDISSEKEYSF 780  
 Qy 781 NPKENKLIVSKYKLHEUSTLLOEIRNNANSDDILEKVMTECEINASNIDRQIVEGR 840  
 Db 781 NPKENKLIVSKYKLHEUSTLLOEIRNNANSDDILEKVMTECEINASNIDRQIVEGR 840  
 Qy 841 TEEPAKNUTSDSINYIKEFKLIESISDSLYDLKHHONGLDDSHFISPDISTENGFRIF 900  
 Db 841 TEEPAKNUTSDSINYIKEFKLIESISDSLYDLKHHONGLDDSHFISPDISTENGFRIF 900  
 Qy 901 INKETGNISIFITEKEFSEYATHISKEISNIKDTIDFNVNGKLVKVNLDAAHEYNTLN 960  
 Db 901 INKETGNISIFITEKEFSEYATHISKEISNIKDTIDFNVNGKLVKVNLDAAHEYNTLN 960  
 Qy 961 SAFFIQSLIENNTKESLSNLNSVAMKRVQYVQLFSTGLNTTDASKVELVSTALDETD 1020  
 Db 961 SAFFIQSLIENNTKESLSNLNSVAMKRVQYVQLFSTGLNTTDASKVELVSTALDETD 1020

## RESULT 2

S10317

toxin B - Clostridium difficile

C;Species: Clostridium difficile

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 15-Oct-1999

C;Accession: S10317; S22434; S22434

R;Barroso, L.A.; Ward, S.Z.; Phelps, C.J.; Johnson, J.L.; Wilkins, T.D.

A;Title: Nucleotide sequence of Clostridium difficile toxin B gene.

A;Reference number: S10317; MNUID:9026540

A;Accession: S10317

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-2366 &lt;BAR&gt;

A;Cross-references: EMBL:X53138; PID:g40442; PID:CAA37298.1; PID:g40443

R;Eichel-Streiber, C.; Laufenberg-Feldmann, R.; Sartingen, S.; Schulze, J.; Sauerborn, M.

Submitted to the EMBL Data Library, July 1991

A;Description: Comparative analysis of Clostridium difficile toxins A and B.

A;Reference number: S21894

A;Accession: S21894

A;Molecule type: DNA

A;Residues: 1-2366 &lt;ECI&gt;

A;Cross-references: EMBL:X60984; PID:g40445; PID:CAA43299.1; PID:g40446

R;von Eichel-Streiber, C.; Laufenberg-Feldmann, R.; Sartingen, S.; Schulze, J.; Sauerborn, M.; Gen. Genet. 233, 260-268, 1992

A;Title: Comparative sequence analysis of the Clostridium difficile toxins A and B.

A;Reference number: S22434; MNUID:92293124

A;Accession: S22434

A;Molecule type: DNA

A;Residues: 1-791-2366 &lt;VOND&gt;

A;Cross-references: EMBL:X60984

C;Genetics:

A;Gene: toxin

C;Superfamily: cpl repeat homology  
C;Keywords: cytotoxin

Query Match 78.7%; Score 4086; DB 2; Length 2366;  
 Best Local Similarity 77.5%; Pred. No. 4.9e-146;  
 Matches 790; Conservative 105; Mismatches 125; Indels 0; Gaps 0;  
 Qy 1 MNLVNAQLQKMYVVKRIOQDEYAILNALBEYHMSSESSVEYLKDINNLTDYVL 60  
 Db 1 MSLVNRKOLEMANVFRQDEYAILDALLEYHMSSESSVEYLKDINNLTDYVL 60  
 Qy 61 NTYKSGRNALKKFKEYLTMEVLELKNNNSLTPVEKLNHFIIWIGGOINDTAINYINQWKD 120  
 Db 61 DTYKSGRNALKKFKEYLTVELELKNNNLTPVEKLNHFIVWIGGOINDTAINYINQWKD 120  
 Qy 121 VNSDVTYKFYYDSDNAFLINTLKKTIVESPATNTLLESPRENLNDPFDYDNKFYKRFMETIY 180  
 Db 121 VNSDYNVNVFDSNAFLINTLKKTVESA-NDTLESFRENLNDPFDYDNKFYKRFMETIY 180  
 Qy 181 DKQKHPTDYKKSQIEBNPEFIDNLTLYSNEKSRLDEALKYIIESLNKNTANNGNDI 240  
 Db 181 DKQKNFTNTYKAQRFENPELJDDIVKTYLIESLNKNTQNSGNDI 240  
 Qy 241 RNLEKPADEDIYRLYQELVERWNLAAASDLRISMLKEDGGYIYDVIDLPGIOPDLPKS 300  
 Db 241 RNFEERKNGESFNLYQELVERWNLAAASDLRISALKEIGGYMDVDMLPGIOPDLPKS 300  
 Qy 301 INKPSTITNTSWEMIKLEAMKYEYIYCYSITSKNFDMLDEEVQSFESALSSKSDKSEIF 360  
 Db 301 IEPKPSVYDWEWTMKLEAMKYEYIPEYTSFEDMDVEQSYFESVFLASKSDKSEIF 360  
 Qy 361 LPDDIKVSPLEVKTIAFANNSVINOLISLKDVSCLDVLVINKRNYKLINDNINPSINE 420  
 Db 361 SSSLGMEASPLEVKIAFNSGTINGLQGLSYKDVSTCNLIVKQIENRYKLINNSUNPASE 420  
 Qy 421 GTDFNTTMKFSKDLIASNEDNMNMKMTNTLVKGFAPDVRSSTINLSPGVTYTGATQD 480  
 Db 421 DNDFNTNTNTFDSIMSAEANDGRFMELGKTDVGFPPDKVITNLSPGEAAAQD 480  
 Qy 481 LLMFDNSTNHLLEPELRNFEPPTKTSQSLTQEITISWSENQARASQOEPEYKKYFE 540  
 Db 481 LLMFKEGSANHILHEADLNFEIRTKNSQSTEDEMASLWSFDARAQFEEYKRNFE 540  
 Qy 541 GALGDDDNLDFAQNTVLKDQYDVSKKLSSMMKTRNKAYHYIVQLOGDKISYEASCNLSK 600  
 Db 541 GSLGBDDDNLDQSNTIVVYDKELEKISSLARSERYHYIVQLOGDKISYEACNLFEK 600  
 Qy 601 DPYSSILOKNIESETAYYYVADAETKEIDYKRIPIQISNRNKLTFIGHGKSEFT 660  
 Db 601 TPYDSVLFQKNIDSEIAYYNNPGDGEIQEIDYKPIISDRPKKTFIGHGKDEFT 660  
 Qy 661 DTFANLVDYDLSLSSSTETLNLAKADISPKYIEINLGCMNPMSYISIAYETPGKLLIK 720  
 Db 661 DIFAGFDDSLSTEALDIAKEDDSKTSIEINLGCMNPMSYISIAYETPGKLLIK 720  
 Qy 721 DRYSELMPSSISODSISITYSANQEVIRINEBKGREILDSGKWNKNEESIKDISKEYSF 780  
 Db 721 DKISELMPSSISODSISITYSANQEVIRINEBKGREILDSGKWNKNEESIKDISKEYSF 780  
 Qy 781 NPKENKIVTKSKYLTLESTLQEBIRNNANSSDILEKRMVLTCEINASNDTQIVER 840  
 Db 781 NPKENKIVTKSKYLTLESTLQEBIRNNANSSDILEKRMVLTCEINASNDTQIVER 840  
 Qy 841 IEPAKNLTSDSINYIKRFKLEISIPSYDLYDKHONGLDDSHFISDTSKTEENGPRIF 900  
 Db 841 IEPAKNLTSDSINYIKRFKLEISIPSYDLYDKHONGLDDSHFISDTSKTEENGPRIF 900  
 Qy 901 INKETGNISIFIEKEFSEYATHISKEISNIKDTIDFNVGKLVKVNLDAAHEYTNL 960  
 Db 901 INKETGNISIFIEKEFSEYATHISKEISNIKDTIDFNVGKLVKVNLDAAHEYTNL 960  
 Qy 961 SAFFIQSLIENNTKESLSNLNSVAMKRVQYVQLFSTGLNTTDASKVELVSTALDETD 1020

Sat Au.